" تحليل قوة الاختبار الإحصائى وعلاقتها بمستوى الدلالة وحجم التأثير فى البحوث التربوية "

د/ محمد إبراهيم محمد محمد

- **مستخلص الدراسة** :

هدفت الدراسة إلى تحليل قوة الاختبار الإحصائى وعلاقتها بمستوى الدلالة وحجم التأثير. تكونت عينة الدراسة من (٩٠) بحث منشور بمجلة البحث فى التربية وعلم النفس للأعداد الصادرة من (٢٠١٠) إلى (٢٠١٢) ، بكلية التربية ، جامعة المنيا ، وأوضحت النتائج أن (٧٢٪) من الفروض تم رفضها ، ومتوسط القوة الإحصائية لها(٠,٢٢) ، و(٢٨٪) من الفروض قبلت، ومتوسط قوة الاختبار (٠,٩٢) ، وجاءت نسب استخدام حجم التأثير مع مستوى الدلالة لاختبار "ت" (٢٠٪) ، وتحليل التباين(٥٪) ، ومعامل الارتباط(صفر) ، والإحصاء اللابارامترى(٣,٣٪) ، وتحليل الانحدار(١,٦٪) ، وبلغت مستويات حجم التأثير للفروض الدالة باختبار "ت"(١٧٪)ضعيف ،(١١٪)متوسط ،(٧٢٪) كبير بينما الفروض غير الدالة (٩٥٪)ضعيف ، (١,٢)متوسط ، (٣,٦٪) كبير ، وبالنسبة لتحليل التباين للفروض الدالة(٧٨٪)ضعيف ،(١١٪)متوسط ،(١١٪) كبير ، والفروض غير الدالة (١٠٠٪) ضعيف كما اتضح وجود علاقة ارتباطية طردية دالة بين قوة الاختبار الإحصائى وكل من حجم التأثير ومستوى الدلالة .

Abstract

This study aimed at analyzing of the statistical power and its relation to the significance level and the effect size.the sample of the study consisted of (90)published researches in the Journal of Research in Education and Psychology from (2010) to(2012). The results showed that (72%)of the hypotheses had been rejected and the means of their statistical power was(0.22),the other (28%) of the hypotheses had been accepted and the means of their statistical power was(0.92), the ratio of using the effect size with the significance level for "t-test" was(20%), the analysis of variance was(5%),the correlation coefficient was(zero),the statistical of Non-parameter was(3.3%), and the analysis of regression was(1.6%). The levels of the effect size for the significant hypotheses using "t-test" were (17%)low,(11%)medium,and (72%)high where the levels of the effect size for the insignificant hypotheses using "t-test" were (95%)low ,(1.2%) medium,and (3.6%)high , the analysis of variance for the significant hypotheses was (78%)low,(11%)medium,(11%)high whereas the analysis of variance for the insignificant hypotheses was(100%) weak. The results also showed a significant correlating relation between the statistical power and the both of the effect size and the significance level.

- **مقدمة** :

اهتم الباحثون في مجال علم النفس والتربية بالتحقق من دقة النتائج التى يتم التوصل إليها عند التحقق من صحة الفروض ، لذلك اعتمدوا على عدة مفاهيم هى : الدلالة الإحصائية ، والدلالة العملية التى يطلق عليها حجم التأثير، و مفهوم قوة الاختبار.

فعند فحص الفرض الصفرى يحاول الباحث تجنب الوقوع فى الخطأ ، والذى يصنف إلى نوعين: الخطأ من النوع الأول ويتمثل فى رفض الفرض الصفرى وهو

فى واقع الأمر صحيح ، والخطأ من النوع الثانى والذى يتمثل فى قبول الفرض الصفرى أو الفشل فى رفض الفرض الصفرى وهو فى الحقيقة خطأ . (Cohen,1992)

ويطلق على الخطأ من النوع الأول " خطأ الفا " حيث تكون القيمة القصوى لاحتمال الوقوع فى مثل هذا الخطأ مساوية لقيمة ألفا والتى يطلق عليها مستوى الدلالة والتى تحدد بمستويين هما (٠,٠٥ أو ٠,٠١)،أما الخطأ من النوع الثانى فيطلق عليه " خطأ بيتا" حيث يقبل الباحث بالفرض الصفرى رغم أنه فى الأساس خطأ ، والفرض البديل هو الصحيح ، وقد يرجع ذلك لعدم توافر أدلة كافية لدى الباحث أو لصغر حجم العينة ، ويشير الباحثون أن قيمة بيتا لاتزيد عن (٠,٢٠) . (Cohen,1994)

ولتجنب الخطأ من النوع الأول يستخدم الباحث الدلالة الإحصائية التى تساعده فى التمييز بين الفرق الحقيقى والفرق الذى يرجع إلى الصدفة أو خطأ المعاينة ، إلا أنها قد تظهر الفرق الكبير الملحوظ على أنه غير حقيقى بين مجتمعين ، وذلك لصغر حجم العينة بينما تظهر الفرق الصغير بين مجتمعين على أنه حقيقى ودال إحصائيا عندما يكون حجم العينة كبيرا، وقد يكون الفرق الكبير فى العينة الصغيرة حقيقى ، وبالتالى فنتيجة الدلالة الإحصائية ترجع لكبر أو صغر حجم العينة (Oleinik&Algina,2000).

لذا فإن مستوى الدلالة له أهمية كبيرة فى البحث حيث يتوقف عليه رفض الفرض الصفرى أو عدم رفضه ، وبالتالى تأييد فرض البحث أو عدم تأييده ، لذا ينبغى عند تحديده أو تفسيره أن يراعى الباحث عدة أمور تتلخص فى تقدير مدى الخطأ أو المخاطرة التى يكون الباحث مستعدا لاحتماله ، ويرتبط ذلك بدرجة خطورة النتائج المترتبة على القرار بقبول أو رفض الفرض الصفرى . مع مراعاة تأثير حجم العينة على النتائج ، فإذا كان حجم العينة كبيرا فينبغى اختيار مستوى دلالة أكثر تشددا ، ولذا يجب أن يتخذ الباحث قراره بتحديد مستوى الدلالة عند مرحلة وضع خطة البحث قبل البدء فى جمع البيانات . (مصطفى حسين باهى ،٢٠١١)

ولكى يتجنب الباحث الوقوع فى الخطأ من النوع الثانى ظهر مفهوم قوة الاختبار الإحصائي ليعبر عن احتمالية الكشف عن حجم التأثير غير الصفرى وبمعنى آخر هو احتمالية عدم الوقوع فى الخطأ من النوع الثانى فعندما يكون الفرض الصفرى خطأ ولكن القيمة المحسوبة كانت دالة عند مستوى (٠,٠٥) فيضطر الباحث لقبول الفرض الصفرى وهو خطأ ، عند ذلك يشير الباحث لعدم توافر الأدلة الكافية لرفض الفرض الصفرى أو الفشل فى رفض الفرض الصفرى . (كما يتضح من الجدول ١)

جدول(١) : توضيح الخطأ من النوع الأول والخطأ من النوع الثانى

العينة عند مستوى دلالة (٠,٠٥)	الفرض الصفرى صحيح حجم التأثير=٠,٠٠	الفرض الصفرى خطأ حجم التأثير≠٠,٠٠
غير دال عدم رفض الفرض الصفرى	لا يوجد خطأ صحيح عدم رفض الفرض الصفرى	خطأ من النوع الثانى(β) خطأ عدم رفض الفرض الصفرى
دال رفض الفرض الصفرى	خطأ من النوع الأول (α) خطأ رفض الفرض الصفرى	لا يوجد خطأ صحيح رفض الفرض الصفرى

يتضح من الجدول(١) تناول ثلاثة مكونات هى (الدلالة الإحصائية ، وحجم التأثير ، ونوع الخطأ)، فعند البحث عن مستوى دلالة (٠,٠٥) فإن قوة الاختبار تحاول الكشف عن حجم التأثير غير الصفرى ، حيث يفيد مستوى الدلالة فى تجنب الخطأ من النوع الأول بينما قوة الاحتبار تفيد فى تجنب الخطأ من النوع الثانى والذى يتمثل فى قيمة بيتا(β) والتى تقدر بأربعة أمثال مستوى الدلالة (٠,٠٥) ، ومن ثم إذا كانت قيمة بيتا لا تزيد عن (٠,٢٠) فإنه يمكن التعبير عن قوة الاختبار الإحصائي بالمعادلة الرياضية التالية : Power=1-B

وبالتالى نستنتج أن أفضل قيمة لقوة الاختبار الإحصائى تكون (٠,٨٠) . (Hallahan &Rosenthal,1996) (Cohen,1992)

و كما يتضح من الجدول (١) فإن حجم التأثير يشير إلي درجة وجود الظاهرة في المجتمع أو درجة الانحراف عن الفرض الصفري فإذا كان الفرض الصفري خاطئ فهو خاطئ لدرجة ما ، وحجم التأثير هو قيمة غير صفرية في المجتمع تزداد مع زيادة درجة الانحراف عن الفرض الصفري ، وبالتالى إذا كان فإذا كان مستوى الدلالة يوضح وجود فروق ترجع للصدفة أم لا ؟ ، فإن حجم التأثير يوضح حجم هذه الفروق فى حال الفروض الفارقة أو قوة العلاقة فى حال الفروض الارتباطية . (Nix and Barnette ,1998)

وبالتالي فهو يفيد في تفسير نتائج الاختبارات الإحصائية بشكل مناسب مستقلا عن تأثير حجم العينة فهو يحدد المقدار الفعلي للفروق بين متوسطات المجموعات أو درجة التلازم بينها، حيث يشير إلي مدى قوة أو ضعف الفرض الصفري فعندما تتفق البيانات التي تم الحصول عليها من التجريب مع الفرض الصفري تماما تصبح القيمة المناظرة لحجم التأثير تساوى صفر ، وتبتعد قيمة حجم التأثير عن الصفر عندما ينتج عن التجربة بيانات غير متوافقة مع الفرض الصفري. (Synder ,Lawson,1993, Kellow,1998, Mcclain,1995)

وتتضح أهمية حجم التأثير فيما أشارت إليه دراسة (Parker & Hagan,2007) لبعض مميزات استخدام حجم التأثير ، حيث يعد طريقة فعالة لتوثيق النتائج فهو مؤشر له تدريج مستمر يدعم القرارات التدريجية فى المعالجة الإحصائية ، ويتميز بالدقة خاصة في حالة النتائج الصغيرة وغير الواضحة ، ويتميز بأنه مقياس جيد لاستقلال النتائج ، خصوصا عند استخدام فترات الثقة ، وفى حالة الصدق وقوة التصميم التجريبي يستخدم كمقياس موضوعى لقوة التدخل ويعد طريقة للمقارنة بين النجاحات المختلفة للتدخل عبر أبحاث دراسة الحالة ، فهو مقياس معتمد فى أبحاث دراسة الحالة المقارنة مع التقاليد البحثية الأخرى، كما أنه يقدم تقديرا لدرجة وجود الظاهرة فى المجتمع ، وذلك من خلال قيمة على مقياس متصل يبدأ من صفر حقيقى (يعنى غياب الظاهرة)، ويمتد ليقدم قيمة لحجم التأثير لهذه الظاهرة .

لذا ينصح الباحثون في مجال العلوم السلوكية ضرورة تقرير حجم التأثير في نتائج دراساتهم إلي جانب اختبار الدلالة الإحصائية ، حيث قد يقع الباحث في خطأ رفض الفرض الصفرى بينما هو صحيح أي أن الفرق بين المتوسطين أو الحدثين صغير أو غير منطقي ، لذا يتحقق الباحث من أن الفرض الصفرى غير

متحقق عندما يكون الفرق الحقيقي يساوى صفرا أو أن الفرق ضئيل للغاية لدرجة أنه يمكن إهماله .(Cashen &Geiger, 2004)

ومن ثم فإن حجم التأثير يشير إلى قوة العلاقة بين المتغيرين في حال الفروض الارتباطية وحجم التأثير الحقيقي في حال الفروض الفارقة والتأثيرية لذا يمكن تصنيف مقاييس حجم التأثير إلى نوعين هما :

- مقاييس الفروق : والتى تدل على مقدار الأثر الذى يحدثه المتغير المستقل فى المتغير التابع ، وتعرف هذه المؤشرات بأنها مقاييس الفرق بين المتوسطات حيث أن هذا الاجراء يجيب عن التساؤل الذى يبحث عن مقدار الفرق الموجود بين المجموعات المبنى على وحدة قياس المتغير التابع، وتتصف بأن لها بنية مشتركة تتمثل فى التغير نتيجة الفروض الخاصة بالتباين فى ضوء الخطأ المعيارى ومن أمثلتها مؤشر كوهين (Cohen,1988)، مؤشر جلاس (Glass,1976)، مؤشر هدجز (Hedges,1981) .
- مقاييس تفسير التباين : وهى تلك المؤشرات التى تدل على قوة العلاقة أو الارتباط بين متغيرات الدراسة ، ويستخدم فى الدراسات الارتباطية والتنبؤيه ولها بنية مشتركة تتمثل فى التغيرات الناتجة عن معادلة مربع معامل الارتباط التى يشار إليها بمصطلح نسبة التباين المفسر ومن أمثلتها : مؤشر مربع إيتا لبيرسون (Pearson,1905)، مؤشر مربع أوميجا لهاس (Hays,1963)، ومؤشر مربع ابيلسون لكيلي (Kelley,1935) (فى : السيد محمد أبو هاشم ، ٢٠٠٥، ويحى نصار ،٢٠٠٦)

ويتضح من التصنيف السابق تعدد المؤشرات الإحصائية المستخدمة فى حساب حجم التأثير (جدول ٢)، نظرا لعناية الباحثين في مجال الإحصاء التربوى والنفسى بعملية اتخاذ القرار الناتجة عن نتائج تلك البحوث والدراسات ، حيث اتجهوا لصياغة مؤشرات إحصائية خاصة بحساب الدلالة العملية لتعزيز القرار الإحصائى (ابتسام حسن الصائغ ، ١٩٩٦)

جدول (٢) : مؤشرات حجم التأثير

المؤشر	الرمز	المعادلة	اختبار الدلالة الإحصائية المناظر	التفسير للقيمة	
d لكوهين	d	$\frac{\bar{x}_{group1}-\bar{x}_{group2}}{SD_{pooled}}$	الفرق بين متوسطين للعينات المستقلة	٠,٢٠	صغير
				٠,٥٠	متوسط
				٠,٨٠	كبير
إيتا تربيع	η^2	$\frac{t^2}{t^2+(n_1+n_2-1)}$	الفرق بين متوسطين للعينات المستقلة	٠,٠١	صغير
				٠,٠٦	متوسط
				٠,١٤	كبير
إيتا تربيع	η^2	$\frac{t^2}{t^2+(n-1)}$	الفرق بين متوسطين للعينات غير المستقلة	٠,٠١	صغير
				٠,٠٦	متوسط
				٠,١٤	كبير

المؤشر	الرمز	المعادلة	اختبار الدلالة الإحصائية المناظر	التفسير للقيمة	
جلاس دلتا	Δ	$\frac{\bar{x}_{Experimental}-\bar{x}_{Control}}{SD_{Control}}$	التحليل التعريفي	٠,٢٠ ٠,٥٠ ٠,٨٠	صغير متوسط كبير
g لهيدجز	g	$\frac{\bar{x}_A-\bar{x}_B}{\sqrt{MS_w}}$	الفرق بين متوسطين للعينات المستقلة	٠,٢٠ ٠,٥٠ ٠,٨٠	صغير متوسط كبير
معامل ارتباط بيرسون	r	$\sqrt{\frac{t^2}{t^2+df_{within}}}$	معاملات الارتباط الخطي والارتباط الخطي الجزئي	٠,١ ٠,٣ ٠,٥	صغير متوسط كبير
معامل التقدير	R^2	$\frac{SS_{regression}}{SS_{total}}$	تحليل الانحدار المتعدد	٠,٠١ ٠,٠٦ ٠,١٤	صغير متوسط كبير
إيتا تربيع لمتغير واحد	η^2	$\frac{SS_{between}}{SS_{total}}$	تحليل التباين تحليل التغاير تحليل الانحدار الخطي	٠,٠١ ٠,٠٦ ٠,١٤	صغير متوسط كبير
إيتا تربيع متعددة المتغيرات	η^2	$1-\Lambda^{\frac{1}{s}}$	تحليل التباين المتعدد تحليل التغاير المتعدد تحليل الانحدار المتعدد	٠,٠١ ٠,٠٦ ٠,١٤	صغير متوسط كبير
أوميجا	ω	$\sqrt{\frac{4\chi^2}{n-\chi^2}}$	مربع كاي χ^2 لجودة التوفيق مربع كاي χ^2 للاستقلال	٠,١ ٠,٣ ٠,٥	صغير متوسط كبير
مربع معامل الارتباط المعدل	R^2_{adj}	$R(1-R^2)\frac{(N-1)}{(N-K-1)}$	تحليل الانحدار المتعدد	٠,٠١ ٠,٠٦ ٠,١٤	صغير متوسط كبير
إيتا تربيع الجزئية	η^2	$\frac{SS_{between}}{SS_{between}+SS_{error}}$	تحليل التباين تحليل التغاير	٠,٠١ ٠,٠٦ ٠,١٤	صغير متوسط كبير
هايز أوميجا تربيع	ω^2	$\frac{SS_{between}-(v-1)MS_{within}}{SS_{total}+MS_{within}}$	تحليل التباين تحليل التغاير	٠,٠١ ٠,٠٦ ٠,١٤	صغير متوسط كبير

المؤشر	الرمز	المعادلة	اختبار الدلالة الإحصائية المناظر	التفسير للقيمة	
مربع إبسيلون	ϵ^2	$\frac{SS_{between}-(\nu-1)MS_{within}}{SS_{total}}$	تحليل التباين تحليل التغاير	٠,٠١	صغير
				٠,٠٦	متوسط
				٠,١٤	كبير
هيرزبرج	R^2	$1-(1-R^2)\left[\frac{(n-1)}{(n-k-1)}\right]\left[\frac{n+k+1}{n}\right]$	تحليل التباين تحليل التغاير	٠,٠١	صغير
				٠,٠٦	متوسط
				٠,١٤	كبير

$\bar{X}$ = المتوسط الحسابي $SS_{between}$ = مجموع المربعات بين المجموعات S الانحراف
ا... $df\,MS_{wn}$ = متوسط ... المجموعات SD_{pooled} = الانحراف المعياري
المجمع ... $SS_{regression}$ = درجات الحرية داخل المجموعات t = قيمة ت المحسوبة في
اختبارت = مجموع مربعات الانحدار

(فى : محمد إبراهيم أحمد الشاردى ، ١٤٣٣هـ)

ونظرا لتعدد مؤشرات حجم التأثير فقد تعددت مؤشرات الحكم على قيمة حجم التأثير من حيث المقدار وكيفية التفسير حسب نوع حجم التأثير وقد اشار كوهين إلى مقدار حجم التأثير فى جداول كوهين لحجم التأثير حيث أوضح العلاقة بين مستويات التأثير فى كل من مربع إيتا ومؤشر كوهين ومربع معامل الارتباط والجدول التالى يوضح ذلك .

جدول (٣) : مستويات حجم التأثير

م	نوع المقياس	مستويات حجم التأثير		
		كبير	متوسط	صغير
١	مربع إيتا	٠,١٤	٠,٠٦	٠,٠١
٢	مؤشر كوهين	٠,٨	٠,٥	٠,٢
٣	معامل الارتباط الثنائى	٠,٣٧	٠,٢٤	٠,١

أشار كوهين إلى أن حجم التأثير(٠,٢٠) تكون درجة تطابق المجتمع الأول والمجتمع الثانى كبيرة، حيث تصل إلى (٨٥٪)، مما يدل على التقارب وانخفاض حجم التأثير ، لذا عندما تبلغ قيمة مؤشر كوهين(٠,٢) فإن حجم التأثير يكون صغير أو ضعيف ، بينما تقل درجة التطابق بين المجتمعين عندما تصل القيمة إلى (٠,٥)، حيث تصل إلى (٦٧٪)، وهو يدل على حجم تأثير متوسط للمعالجة الإحصائية فى حالة اختبار "ت" ، وتقل درجة المطابقة بين المجتمعين عندما تصل قيمة حجم التأثير إلى (٠,٨٠) ، حيث تصل المطابقة إلي (٥٣٪) وهذا يدل على حجم تأثير كبير .(أحمد سليمان عوده، و خليل يوسف الخليلي، ٢٠٠٠ (Howell,1995

كما يلاحظ من الجدول أنه عندما تكون قيمة مؤشر كوهين (٠,٢)، فإن ذلك يعنى أنها تمثل(٧,٩٣٪) من مساحة المنحنى الاعتدالى فى أحد طرفيه وبالتالى فإن حجم التأثير يعد صغيرا، وعندما تصل إلى (٠,٥) فإنها تمثل (١٩,١٥٪) من مساحة المنحنى الاعتدالى عند أحد طرفيه وبالتالى يكون حجم

التأثير متوسطا ، وعندما تبلغ (٠,٨) فإن هذا يعنى أنها تمثل (٢٨,٨١٪) من مساحة المنحنى فى أحد طرفيه .(خليل أحمد محمود لبد، ٢٠٠٥)

والبحث الحالى يهدف لدراسة العلاقة بين قوة الاختبار الإحصائي وكل من الدلالة الإحصائية وحجم التأثير، وذلك من خلال رصد واقع استخدام قوة الاختبار الإحصائي فى البحوث التربوية المنشورة منذ (٢٠١٠) وحتى (٢٠١٢)، ثم دراسة العلاقة بين تلك المفاهيم فى تلك البحوث المنشورة .

- **مشكلة البحث** :

تعد الإحصاء من الوسائل المهمة فى مجال البحوث التربوية والنفسية حيث تفيد فى إيضاح النتائج وتجعلها أكثر إقناعا وذات معنى للقارئ كما أنها توضح حقيقة الظاهرة وحقيقة الفروق بين المجتمعات ومن ثم فهى ضرورة للبحث العلمى والتربوى ولكنها وسيلة وليست غاية لذا يتحدد استخدامها بمدلول معناها ومدى مناسبة الأسلوب المستخدم للبيانات المشتقة من المجتمع والفروض البحثية وامكانية تفسيرها تفسيرا ذو معنى واضح للقارئ ، وبالتالى فإن الباحث يجب أن يحدد بدقة الأسلوب الإحصائى المناسب لطبيعة بياناته وعينته وفروضه وما يهدف إليه .

ويعتمد البحث التربوى والنفسى اختبارات الدلالة الإحصائية كأسلوب علمى لتقرير ما إذا كانت الفروق أو الارتباطات أو العلاقات بين المتغيرات ذات دلالة حقيقية أم ترجع للصدفة ، فالدلالة الإحصائية شرط ضرورى لصناعة قرار تربوى أو نفسى ولكنه ليس كافيا حيث تحول استخدامها فى البحوث من أساليب إلى طقوس تفقدها معناها. (Cohen,1977)

لذا برزت الدلالة العملية نتيجة لسوء استخدام اختبارات الدلالة الإحصائية حيث يعتمد الباحثون فى تقدير نتائجهم على الدلالة الإحصائية دون محاولة الكشف عن مقدار العلاقة القائمة بين المتغيرين رغم عدم اهميتها من الناحية التطبيقية أو العملية فإذا وجد الباحث أن القيمة دالة إحصائيا فهذا يعنى أن المتغير المستقل له تأثيره على المتغير التابع، ولكنه لا يدل على حجم التأثير أو درجة العلاقة بين المتغيرين ، وربما كانت الدلالة الإحصائية هنا لا تعنى وجود علاقة قوية بين المتغيرين (زكريا الشربينى، ٢٠٠٧)

حيث أشارت دراسة (Hasse and Waechter and Solonon,1982) إلى واقع الدلالة العملية بواسطة مربع إيتا للاختبار الإحصائى (F) المستخدمة فى مجلة علم النفس الارشادى للمجلدات (١٧- ٢٦) وللفترات ما بين(١٩٧٠ -١٩٧٩) وتكونت عينة الدراسة من (١٠٧) مقال بواقع (١٠٤٤) اختبارا إحصائيا حيث أكدت النتائج أن حجم التأثير محسوبا بمربع إيتا للإحصائى (F) كان صغيرا حيث بلغت قيمته (٠,٠٨).

واتفقت مع تلك النتيجة ما توصلت إليه دراسة (عبد العاطي أحمد الصياد ١٩٨٨) التي هدفت للتعرف على واقع الدلالة العملية وحجم العينة المصاحبتين للدلالة الإحصائية لاختبار "ت" فى البحث التربوى والنفسى حيث تكونت العينة من(٢٦) دراسة في مجالى علم النفس والصحة النفسية ومجالات تربوية أخرى

و(١٣) أطروحة فى مجال علم النفس والتربية حيث بلغ تكرار استخدام الاختبار الإحصائى "ت" (٤٦٣ مرة) ، وقد أكدت النتائج أن الدلالة العملية للاختبار الإحصائى "ت" من النوع الضعيف .

وتوصلت دراسة (عبد الله عمر النجار، ١٩٩٠) لتقويم الوضع القائم لاستخدام الأساليب الإحصائية فى رسائل الماجستير فى كل من كليتى التربية بجامعة أم القرى وجامعة الملك سعود حيث تكونت عينة الدراسة من (١٧٧) رسالة ماجستير موزعة كالتالى :(١٢٧) من الرسائل التى قدمت إلى كلية التربية بجامعة أم القرى، و (٥٠) رسالة من الرسائل التى قدمت إلى كلية التربية بجامعة الملك سعود وذلك واقع (٥٠٪) من كل مجتمع من مجتمعى الدراسة ولحساب الدلالة العملية اختار الباحث عشوائيا (٣٥٣) اختبارا إحصائيا من كلا الكليتين ، وأظهرت النتائج أن الدلالة العملية للأساليب الإحصائية الشائعة ضعيفة جدا ، وأن الاستخدام غير المناسب للأساليب الإحصائية تكرر بنسبة (٤٧٪) تقريبا.

وأكدت دراسة (Shaver,1992) أن الدلالة الإحصائية لا تشير إلى احتمالية خطأ أو صحة الفرض الصفرى ، ويجب تشجيع الباحثين على التقليل من الاعتماد على الدلالة الإحصائية فى تحليلاتهم وتفسيراتهم ولا ينبغى نشر الأبحاث التى تعتمد على الدلالة الإحصائية فقط ، ولا بد من احتواء التحليلات ونتائج الدراسات على أحجام التأثير .

لذا فإن اهمال الدلالة العملية يقلل من أهمية نتائج الدراسة وقد يؤدى إلى التضليل لأن الأساس فى البحث التربوى والنفسى هو تقدير الدلالة العملية فالباحثون الذين لا يعتمدون على الدلالة الإحصائية وحدها ويلجأون إلى حساب حجم التأثير وقوة الاختبارات الإحصائية يساهمون فى تحسين جودة البحوث (Huston,1993)

حيث أكدت نتائج دراسة ابتسام حسن الصائغ (١٩٩٦) . التى هدفت لدراسة أهم الطرق الإحصائية لحساب وتقدير الدلالة العملية للاختبار الإحصائى "ت" وتحليل التباين التى استخدمت فى (١٠٣) رسالة ماجستير . إلى أن الدلالة العملية المصاحبة للدلالة الإحصائية كانت عند حجم تأثير كبير لاختبار " ت" حيث بلغ (٠,٨) فأكثر بينما كان حجم التأثير لاختبار تحليل التباين صغيرا حيث بلغ (٠,١) وبصفة عامة كانت الدلالة العملية عند مستوى تأثير منخفض بلغ (٦٤٪) كما أن نسبة (٤٧٪) من الفروق الدالة إحصائيا كانت قيم مربع إيتا ومربع أوميجا فيها منخفضة ، كما أوضحت النتائج أنه ليس هناك أثر لمستوى الدلالة الإحصائية على الدلالة العملية.

وقد أوضحت دراسة تحسين حجيمات ، وخليل عليان (١٩٩٧) واقع الدلالة الإحصائية وعلاقتها بكل من الدلالة العملية وقوة الاختبار حيث تكونت العينة من (١٨٨) فرض صفرى استخدم فى فحصها اختبار تحليل التباين وذلك في (٢٨) رسالة من رسائل الماجستير وقد أكدت النتائج أن حوالى (٨٥٪) من الفروض الصفرية استخدم فى فحصها مستوى الدلالة (٠,٠٥) أما الباقي

فاستخدم في فحصها مستوى الدلالة (٠,٠١) وأن حوالى(٦٣٪) من الفروض التى كانت دالة إحصائيا لم تكن ذات دلالة عملية كما أشارت النتائج أن (٣,٢٪) من الفروض الصفرية المستخدم فى تحقيق منها تحليل التباين كانت غير دالة إحصائيا ولكن لها حجم أثر كبير عمليا .

ومن ثم يتضح للباحث من استعراض الدراسات السابقة عدم وجود علاقة بين الدلالة الإحصائية والدلالة العملية حيث أن نسبة كبيرة من الفروض الدالة إحصائيا جاءت دلالتها العملية ضعيفة ونسبة ليست قليلة من الفروض غير الدالة إحصائيا كانت دلالتها العملية كبيرة مما يستدعى ضرورة ارفاق الدلالة العملية مع الدلالة الإحصائية فى البحث التربوى والنفسى لأن كل منهما مكمل للآخر، كذلك يتضح ضرورة الاعتماد على تحديد قوة الاختبار الإحصائي لأنه يؤثر على كل من الدلالة الإحصائية والدلالة العملية فقد أوضحت دراسة تحسين حجيمات ، و خليل عليان (١٩٩٧) أن نصف الاختبارات الإحصائية المستخدمة فى فحص الفرض الصفرى والدالة إحصائيا تمتعت بقوة اختبار من مستوى متوسط وكبير وأن حوالي (٢٥ ٪) من الاختبارات الإحصائية لم تكن دالة إحصائيا وكانت قوتها ضعيفة .

لذا أشارت دراسة (Mclean & Eronest,1997) إلى ضرورة الاشارة إلى قيم حجم التأثير كنتيجة مكملة لنتائج فحص الفروض وكمعيار آخر لفهم النتائج الإحصائية فإذا كانت الدلالة الإحصائية تفحص احتمال الحصول على نتائج دالة بمحض الصدفة فإن حجم التأثير يفحص مقدار الدلالة العملية للنتائج .

فالدلالة الإحصائية تهتم بمستوى الثقة التى نوليها للنتائج فمثلا الفرق الدال عند مستوى (٠,٠٥)يدل على وجود فروق بين المجموعتين وأن هذه الفروق حقيقية لا ترجع للصدفة وأن هذا الحكم نثق فيه بنسبة (٩٥٪) ولكن يؤخذ على الدلالة الإحصائية أنها تتأثر بحجم العينة فإذا زاد حجم العينة وصل الفرق لمستوى الدلالة ولو كان الفرق صغيرا، وإذا قل حجم العينة لم يصل الفرق لمستوى الدلالة ولو كان الفرق كبيرا، بينما تعتمد فكرة الدلالة العملية على صياغة الفروق بين المتوسطات أو التعبير عن العلاقة بين المتغيرات المستقلة من جهة والمتغيرات التابعة من جهة أخرى عن طريق استخراج حجم تباين المتغير التابع الذى يمكن تفسيره عن طريق المتغير المستقل كما هو الحال فى تحليل التباين (رشدى فام منصور، ١٩٩٧، Cohen ,1988)

وهدفت دراسة ألطاف إبراهيم (١٩٩٨) إلى التعرف على واقع الدلالة الإحصائية والعملية وقوة الاختبار فى رسائل الماجستير فى الجامعة الاردنية حيث اشتملت الدراسة على (٣١٠) فرض صفرى استخدم فى فحصها اختبار تحليل التباين ومعامل الارتباط ومربع كاى واختبار "ت"، وأظهرت النتائج أن حوالى (٣٠,٤٪) من الفروض الدالة إحصائيا كانت دلالتها العملية معدومة تقريبا ، وأن حوالى (٣٠,٩ ٪) من الفروض الدالة إحصائيا كانت دلالتها العملية ضعيفة وبالتالى فإن حوالى (٦١٪) من الفروض الدالة إحصائيا كانت دلالتها

العملية معدومة أو ضعيفة وأن حوالى (١١,٨٪) من الفروض الدالة إحصائيا تمتعت بدلالة عملية من النوع المتوسط والكبير ، بينما بلغت نسبة (١٠,٤٪) من الفروض غير الدالة إحصائيا بدلالات عملية متوسطة وكبيرة ، وأن حوالى (٥٦٪) من الفروض غير الدالة إحصائيا كانت قوة اختباراتها الإحصائية ضعيفة ، وكانت قيم حجم التأثير لتحليل التباين (٠,٢٥) ومعامل الارتباط (٠,١) ومربع كاى (٠,٣٩) بينما اختبار "ت" (٠,١٦) وهذا يؤكد أن حجم التأثير لكل من معامل الارتباط واختبار "ت" كانت ضعيفة بينما تشير نتائج اختبار تحليل التباين ومربع كاى إلى أن قيم حجم التأثير متوسطة طبقا لتصنيف كوهين.

وتوصلت دراسة إبراهيم رشاد (٢٠٠٠) ـ لواقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار للاختبارات الإحصائية المستخدمة فى بحوث مجلة دراسات العلوم التربوية فى الجامعة الاردنية . أن حوالى(٩٥٪) من الفروض استخدمت مستوى الدلالة (٠,٠٥) وأن الدلالة العملية حسبت بالمجلة بنسبة ضعيفة بلغت (٥,٦٪) .

وتوصلت دراسة (Smithson,2001) لأحجام التأثير من خلال مربع معامل الارتباط والذى يستخدم فى تحليل الانحدار المتعدد وبخاصة فى التوزيعات اللامركزية مثل قيمة ف للتباين ومربع كاى إلى ضرورة وضع قيمة (R) فى الاعتبار عند استخدام التنبؤات من خلال معاملات الانحدار وانتهت إلى أن قيمة حجم التأثير متوسطة بالاضافة إلى وجود علاقة قوية بين حجم التأثير وفترات الثقة باستخدام التوزيعات اللامركزية .

وهدفت دراسة عادل أحمد بابطين (٢٠٠١) إلى الكشف عن أهم مشكلات الدلالة الإحصائية فى البحث التربوى والتعرف على أهم المفاهيم والأساليب الإحصائية التى يمكن أن تقدم حلولا لتلك المشكلات وقد تكونت العينة من (٩٩) رسالة من رسائل الماجستير بكلية التربية بجامعة أم القرى ، وقد توصلت الدراسة إلى أن أهم مشكلات الدلالة الإحصائية فى البحث التربوى استخدامها كتفسير للنتائج وكتفسير لتأثير المعالجة، وأن هناك بعض المفاهيم التى يمكن أن تقدم حلولا لمشكلات الدلالة الإحصائية وهى : تحليل القوة الإحصائية تقدير حدود الثقة، وتقديرات حجم التأثير، وتحليل الإعادة، كما أن من الأساليب الجديدة لحل مشكلات الدلالة الإحصائية هى إعادة صياغة الأسئلة البحثية والفروض بحيث تركز على حجم التأثير بدلا من مجرد وجود أو عدم وجود فروق، وقد كشفت الدراسة أن حوالى (٩٨٪) من رسائل الماجستير بكلية التربية بجامعة أم القرى تعتمد على نتائج اختبارات الدلالة الإحصائية فقط فى تفسير النتائج ، وأن حوالي (٨٥,٩٪) من الرسائل تستخدم نتائج اختبارات الدلالة الإحصائية فى تفسير تأثيرات المعالجات أو العوامل المشتقة ومقاييس حجم التأثير استخدمت فقط فى (٢٪) من المجموع الكلى للرسائل .

وأوصت دراسة (portillo,2001) ـ التى هدفت إلى دراسة اختبارات الدلالة الإحصائية وقياسات حجم التأثير حول (١٤١) بحث كمي في جريدة التعليم الزراعى . بضرورة أن تتضمن الأبحاث الدلالة الإحصائية وقياسات حجم التأثير

وتحليل قوة الاختبار وفترات الثقة واعتماد انفاذ سياسات أكثر صرامة فيما يتعلق باختبارات الدلالة الإحصائية وقياسات حجم التأثير.

وتوصلت دراسة يحي نصار (٢٠٠٢) أن حجم التأثير أقل تأثرا بحجم العينة من الأساليب الإحصائية التى تستخدم لفحص الفروض كما أن نسبة (٦٠٪) من قيم حجم التأثير للأساليب الإحصائية كانت صغيرة ومتوسطة .

وفى دراسة (King,2002) التى هدفت لدراسة الدلالة الإحصائية ضد الدلالة العملية فإنها توصى بتطبيق اختبارات الدلالة الإحصائية بشكل مدروس، وليس بطريقة ميكانيكية وبفهم كامل لقوة كل منهما وضعفها وبيان قياس حجم التأثير مع فترة الثقة المناسبة له ، ومحاولة تقييم الدلالة الإحصائية بدلا من استخدام لفظ الدلالة، وذلك لأن استخدام مصطلحات غير دقيقة يشارك فى سوء الفهم، ولأن نتائج الدلالة الإحصائية قد تكون أو لا تكون ذات دلالة نظريا .

كما أوضحت دراسة فريال محمود الحاج (٢٠٠٣) واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار للاختبارات الإحصائية المستخدمة فى رسائل الماجستير الصادرة عن كلية العلوم التربوية بجامعة مؤته ، حيث أكدت النتائج أن معظم الباحثين يستخدمون مستوى الدلالة (٠,٠٥) وأن نسبة الاختبارات الدالة إحصائيا بلغت (٤٥٪) وكان منها حوالى (٧١٪) اختبار ذو دلالة عملية ضعيفة .

أما دراسة عبد الله عمر النجار (٢٠٠٥) فقد هدفت لتقويم استخدام الباحثين فى الأبحاث الادارية المنشورة فى مجلة جامعة الملك سعود – العلوم الادارية للدلالة العملية فى اختبار مربع كاى وأظهرت النتائج أن نسبة (٣٩,٤٦٪) من الفروض المعتمدة على اختبارات مربع كاى دالة إحصائيا ، وأن نسبة (٥٠,١٩٪) كانت ذات دلالة عملية منخفضة أما الفروض ذات الدلالة العملية المرتفعة فقد بلغت حوالي (١٠,٣٤٪) وأكدت الدراسة أن زيادة حجم العينة قد يكون له تأثير سلبى مكلف دون فائدة تذكر على مستوى الدلالة العملية المتحصل عليها ، لذا فإن التعرف على حجم التأثير المناسب عند إجراء الدراسة سيساعد الباحثين على تخفيض التكاليف لمجمل أبحاثهم ، وكذلك على إجراء البحوث والدراسات باستخدام عينات صغيرة والحصول منها على النتائج المطلوبة .

ويلاحظ الباحث من خلال استعراض الدراسات السابقة أنه رغم أن الدراسات التقويمية للبحوث التربوية والنفسية بدأت منذ فترات بعيدة زمنيا إلا أن الدراسات الحديثة نتائجها لم تختلف كثيرا عن الدراسات البعيدة زمنيا حيث القصور على استخدام الدلالة الإحصائية وعدم تناول الدلالة العملية وقوة الاختبار وخاصة فى البحوث العربية كما أشارت تلك الدراسات لأهمية حجم التأثير وأثره على دقة نتائج القرارات الإحصائية وأوصت بضرورة عدم الاعتماد فقط على الدلالة الإحصائية دون الدلالة العملية حيث أن " الدلالة الإحصائية وحدها ليست كافية لصنع قرار تربوى أو نفسى فهى شرط ضرورى فقط ولكنه ليس كافيا ، والكفاية تتحقق إذا ما حسبنا قوة العلاقة بين المتغير التابع والمتغير المستقل وقوة العلاقة هنا المقصود بها الدلالة العملية التى قد يسميها

الباحث النفسى بالدلالة السيكولوجية ويسميها الباحث التربوى بالدلالة التربوية" (عبدالعاطي أحمد الصياد، ١٩٨٨)

فمن الممكن أن تكون الدلالة الإحصائية مضللة لذا من المهم الرجوع إلى حجم التأثير فى تقويم نتائج أية تجربة حيث تخبرنا أحجام التأثير عن مقدار التأثير فى حين لا تخبرنا الدلالة العملية عن ذلك .(Wright,2003)

ورغم تلك الانتقادات الموجهة نحو الاقتصار على الدلالة الإحصائية فقط فى مجال البحث التربوى والنفسى إلا أن التركيز على استخدامها وسوء الاستخدام يبقيان شائعين فى مجال البحث مع أنه يجب التكامل بينها وبين الدلالة العملية .(Shaver,1992)

ومن ثم تحتاج البحوث التربوية والنفسية إلى مراجعات مستمرة ، وذلك من خلال الدراسات التى تهدف للتحليل البعدى للبحوث من حيث التصميم والإجراءات ، ومعايير الصدق والثبات ، وكذلك الدلالة العملية للنتائج .

كذلك يحتاج البحث التربوى لتحليل القوة الإحصائية لنتائج فروضه حيث يعتبر مؤشر قوة الاختبار الإحصائي موضوع ذو أهمية لأي باحث يستخدم فى دراسة منطق الاستدلال الإحصائي ، فالدراسات التى تكون فيها الاختبارات الإحصائية ذات مستويات منخفضة من قوة الاختبار عادة تؤدى إلى نتائج غير حاسمة على الرغم من إنفاق الباحثين للكثير من الوقت والجهد لجمع بيانات أبحاثهم وتحليلها .(عبد الله عمر النجار ، ٢٠٠٦)

والبحث الحالي يهتم بدراسة واقع قوة الاختبار والدلالة العملية فى الأبحاث التربوية المنشورة فى مجلة البحث فى التربية وعلم النفس الصادرة عن كلية التربية جامعة المنيا حيث لا توجد دراسة تناولت مراجعة وتقويم تلك المجلة رغم عراقة تلك المجلة وذلك بهدف الوقوف على واقع استخدام كل من قوة الاختبار والدلالة العملية فى تلك البحوث .

لذلك يطرح الباحث التساؤلات التالية :

- ما واقع قوة الاختبار فى الأبحاث المنشورة بمجلة التربية وعلم النفس ؟
- ما واقع استخدام حجم التأثير فى الأبحاث المنشورة بمجلة التربية وعلم النفس ؟
- ما واقع حجم التأثير للفروض الدالة وغير الدالة إحصائيا لاختبارى "ت " و "تحليل التباين "؟
- هل توجد علاقة ارتباطية دالة بين حجم التأثير وقوة الاختبار الإحصائي ؟
- هل توجد علاقة ارتباطية دالة بين مستوى الدلالة وكل من قوة الاختبار الإحصائي وحجم التأثير ؟

- **أهداف البحث** :

يهدف البحث الحالى إلى :

- تناول قوة الاختبار الإحصائى تعريفها وواقع استخدامها فى البحث التربوى والنفسى من خلال مراجعة مجلة البحث فى التربية وعلم النفس.

- تناول مفهومى الدلالة العملية والدلالة الإحصائية وواقع استخدامهما فى البحث التربوى والنفسى من خلال مراجعة مجلة البحث فى التربية وعلم النفس.
- دراسة واقع حجم التأثير للفروض الدالة وغير الدالة إحصائيا لاختبارى "ت" و"تحليل التباين.
- الوقوف على العلاقة الارتباطية بين حجم التأثير وقوة الاختبار الإحصائي .
- الوقوف على العلاقة الارتباطية بين مستوى الدلالة وكل من قوة الاختبار الإحصائي وحجم التأثير

- **أهمية البحث** :

تتضح أهمية البحث الحالى من خلال تناوله لمصطلحين مهمين هما : قوة الاختبار الإحصائى والدلالة العملية وذلك لما لمؤشر قوة الاختبار من أهمية فى تحديده فى جميع مراحل البحث (قبل، وأثناء، وبعد) حيث أن التحديد القبلى يركز على تحديد الباحث لقيمة معينة لقوة الاختبار وفى ضوئها يتم استكمال إجراءات البحث والدراسة ، فالتعرف على قوة الاختبار الإحصائى ينتج للباحث الارتقاء ببحثه من خلال زيادة احتمالية الحصول على نتائج ذات دلالة وزيادة كفاءتها مما يمكن الباحث من تفسير أفضل للنتائج المتحصل عليها ، وذلك لأن قوة الاختبار هى المحك الأساسى لاختبار الفروض الإحصائية .(عبد الله عمر النجار ، ٢٠٠٦)

كما أن الدلالة العملية (حجم التأثير) تسهل من التحليل المتلاحق حتى مع دمج تقارير مختلفة فيفيد فى تحليل القوة حيث استخدم كوهين مجموعة من المؤشرات لقياس قوة رفض الفرض الصفرى فى العديد من التصميمات علاوة على أنه يستخدم فى صياغة الأبحاث وهو ما يتضح من خلال دور حجم التأثير فى التحليلات البعدية حيث يتم دمج مجموعة من أحجام التأثير فى صورة تقرير واحد لحجم التأثير الشامل ، لتوضيح توزيع أحجام التأثير وبناء النموذج الاستطلاعى للتنوع فى حجم التأثير ، وفى بعض الحالات يكون حجم التأثير المستخدم فى التحليل البعدى هو ذاته المستخدم فى تحليل القوة ، ونظرا لأن حجم التأثير يستخدم فى تقريرالأبحاث فلابد من تزويد القارئ بمعلومات عن كل من الدلالات الإحصائية ومعلومات كافية عن حجم التأثير (Streinberg&Thissen,2006). (فيصل أحمد العبد القادر،٢٠٠٩)

كما تتضح أهمية البحث الحالى فى تناوله الشامل لكل من قوة الاختبار الإحصائى وحجم التأثير حيث يعرض البحث مجمل ما توصل إليه الباحثون فى مجال حجم التأثير عالميا وعربيا مما يجعل البحث متضمنا لجميع المعادلات لأغلب مؤشرات حجم التأثير لكل من الإحصاء البارامترى والإحصاء اللابارامترى مع كيفية التفسير لتلك المؤشرات .

تناول البحث العلاقة بين مستويات الدلالة الإحصائية والدلالة العملية بمستوياتها باختلاف الأسلوب الإحصائى ليوضح العلاقة بين الدلالة العملية والدلالة الإحصائية فى الأساليب الإحصائية البارامترية واللابارامترية لتقدم

للباحث الإجابة عن التساؤل متى يستخدم حجم التأثير ؟ هل يستخدم مع الفروض الدالة إحصائيا أم مع الفروض غير الدالة إحصائيا أم مع الاثنين؟ كما تقف النتائج حول اختلاف دلالة قوة حجم التأثير باختلاف الأسلوب الإحصائى .

يتناول البحث مراجعة تقييميه لمجلة البحث فى التربية وعلم النفس الصادرة عن كلية التربية جامعة المنيا حيث أن تلك المجلة تصدر منذ الثمانينات ورغم عراقتها إلا أنه لا يوجد اهتمام بمراجعة الأبحاث التى تنشر فيها وقوفا على نقاط القوة والضعف وارشادا للقائمين عليها بما فى المجلة من نقاط يجب الانتباه لها لذا فإن مراجعة استخدام قوة الاختبار وكل من الدلالة الإحصائية والعملية فى المجلة وتوضيح ذلك من خلال نتائج البحث يفيد الناشر فى المجلة والقارئ وهيئة التحرير والمحكمين على السواء فى تقييم المجلة وتطويرها .

- **مصطلحات الدراسة** :
- **قوة الاختبار** Power of The Test :

ويقصد بها قدرة الاختبار على رفض الفرض الصفرى عندما يكون خاطئا (Cohen,1992)

- **حجم التأثير** Effect Size :

يعرف بأنه " قياس درجة الفروق أو الارتباطات بين المتغيرات ويطلق عليه الدلالة العملية "(Cohen,1962)

وهو يمثل " الدرجة التى تمثلها الظاهرة فى المجتمع " (Cohen,1988,9)

وبالتالى فهو " يشير إلى نمطين من المؤشرات شائعة فى الأدبيات هما : قياس حجم التأثير طبقا للفروق بين متوسطات المجموعات ، والثانى قياس قوة الارتباط طبقا للتباين ." (Maxwell&Delaney,1990)

ويطلق عليه الدلالة العملية، وهو عبارة عن قيمة كمية تشير إلى درجة العلاقة بين متغيرات الدراسة ضمن مجتمع محدد مسبقا هو مجتمع الدراسة بغض النظر عن الأسلوب الإحصائى المستخدم للتعبير عن ذلك الأثر (يحي نصار ، ٢٠٠٢) .

والذي يفسر، مقدار التباين فى المتغير التابع الذى يمكن تفسيره من خلال متغير مستقل يرتبط أو يؤثر فيه (فى، عبد الله عمر النجار، ٢٠٠٥)

ومن ثم ، فهو الأساليب التى يتم من خلالها معرفة حجم الفرق أو حجم العلاقة بين متغيرين أو أكثر ويطلق عليه مقاييس قوة التأثير أو قوة الترابط (رشدى فام منصور ، ١٩٩٧)

- **الدلالة الإحصائية** :

ويقصد بها درجة اقتراب قيمة مقاييس العينات من مقاييس المجتمع الأصل ويمكن اكتشافها بواسطة الاختبار الإحصائى وهو مجموعة من القواعد تمكن

الباحث من رفض أو قبول الفرض الصفرى وبموجبه يمكن الحكم على الفرض.(عبد العاطي أحمد الصياد، ١٩٨٨)

- **مستوى الدلالة** Significance Level :

هو احتمالية الوقوع فى الخطأ من النوع الأول ويمثل حد أقصى يرتضيه الباحث للخطأ من النوع الأول .(عبد العاطي أحمد الصياد، ١٩٨٨)

- **منهجية البحث** :

- العينة : تم تحليل نتائج (٩٠) دراسة منشورة بمجلة البحث فى التربية وعلم النفس بكلية التربية جامعة المنيا والتى صدرت من تاريخ (٢٠١٠) حتى نهاية (٢٠١٢)
- الأدوات : استمارة التحليل : تم إعداد استمارة التحليل من خلال تضمينها عدد من الجداول تهدف لتجميع البيانات التى تحقق أهداف البحث الحالى على مستوى الدراسات ، وعلى مستوى الفروض .
- بالنسبة للدراسات : تم إعداد جدول يتكون من ثلاثة عناصر هى الأسلوب الاحصائى المستخدم فى الدراسة ، وهل استخدم الباحث مع ذلك الأسلوب حجم التأثير أم لا ؟ والنسبة المئوية للاستخدام وعدم الاستخدام
- بالنسبة للفروض :تم إعداد جدول يتكون من : رقم الدراسة ، رقم الفرض الأسلوب الإحصائي ، قيمة الإحصائي ، درجة الحرية ، مستوى الدلالة ، حجم التأثير بمربع إيتا ، حجم التأثير بكوهين ، قوة الاختبار الإحصائى .

- **إجراءات التطبيق** :

- تم تجميع جميع أعداد مجلة البحث فى التربية وعلم النفس منذ عام(٢٠١٠) وحتى نهاية (٢٠١٢) .
- تم استبعاد الدراسات التى اعتمدت على التحليل الكيفي فقط .
- اقتصر الباحث فى تحليل واقع قوة الاختبار الإحصائي على الفروض التى استخدمت اختبار "ت" لإنها كانت أكثر تكرارا .
- استخدم الباحث المتوسط الحسابى والانحراف المعياري والحد الأدنى والأعلى والنسبة المئوية لرصد واقع قوة الاختبار الإحصائي للفروض التى رفضت الفرض الصفرى والفروض التى قبلت الفرض الصفرى .
- تم حساب توزيع نسب حجم التأثير للأساليب الإحصائية البارامترية واللابارامترية .
- رصد الباحث قوة حجم التأثير لكل من اختبار "ت" واختبار "تحليل التباين" لكل من الفروض الدالة إحصائيا وغير الدالة إحصائيا .
- استخدم الباحث معامل الارتباط لإيجاد العلاقة بين حجم التأثير وقوة الاختبار الإحصائي للفروض الدالة إحصائيا والفروض غير الدالة إحصائيا .
- استخدم الباحث معامل الارتباط لإيجاد العلاقة بين مستوى الدلالة وكل من قوة الاختبار الإحصائي وحجم التأثير .

- **نتائج البحث** :

السؤال الأول:" ما واقع قوة الاختبار الإحصائي فى مجلة البحث فى التربية وعلم النفس ؟"

لاجابة هذا السؤال قام الباحث بدراسة عدد (٢٢٨) فرضا من الفروض الفارقة التى استخدمت اختبار "ت" نظرا لأنها أكثر الاختبارات الإحصائية تكرارا فى المجلة ، كما أن تلك الفروض قدمت بيانات يمكن من خلالها حساب قوة الاختبار الإحصائي، حيث قدم الباحث معلومات عن حجم العينة والمتوسطات الحسابية والانحراف المعياري مما مكن الباحث من حساب قيمة (d) لكوهين كما وفرت بيانات منها مستوى الدلالة ونوعها، وبالتالي تمكن الباحث من استخراج قيمة قوة الاختبار الإحصائي باستخدام برنامج (G*Power v.3.1) والذى يستخدم فى حساب قوة الاختبار الاحصائى لعدة طرق منها الطريقة البعدية المستخدمة فى البحث الحالى

جدول (٤) : واقع قوة الاختبار الإحصائي فى مجلة البحث فى التربية وعلم النفس

	العدد	متوسط قوة الاختبار	الانحراف المعيارى	أعلى قيمة	أقل قيمة	أقل من ٠,٨٠		أعلى من ٠,٨٠	
						العدد	النسبة	العدد	النسبة
رفض الفرض الصفرى	١٦٥	٠,٩٢	٠,١٧	١,٠٠	٠,٠٨	٢١	٩,٢١	١٤٤	٦٣,١٥
قبول الفرض الصفرى	٦٣	٠,٢٢	٠,١٨	٠,٩٩	٠,٠٥	٦١	٢٦,٧٥	٢	٠,٨٧
ككل	٢٢٨	٠,٧٢	٠,٣٦	١,٠٠	٠,٠٥	٨٢	٣٥,٩٦	١٤٦	٦٤,٠٢

يتضح من الجدول (٤) ما يلى :

- **عدد الفروض التى رفضت الفرض الصفرى بلغت (١٦٥) فرضا بنسبة (٧٢٪) بينما بلغ عدد الفروض التى قبلت الفرض الصفرى أو فشلت فى رفضه (٦٣) بنسبة (٢٨٪).**
- **متوسط قوة الاختبار الإحصائي للفروض التى رفضت الفرض الصفرى بلغ (٠,٩٢) وهو يشير إلى ارتفاع قوة الاختبار الإحصائي لتلك الفروض بينما تراوحت قيمته بين (٠,٠٨ إلى ١,٠٠) ، أما متوسط قوة الاختبار للفروض التى قبلت الفرض الصفرى بلغ (٠,٢٢) وهى قيمة أقل من (٠,٨٠) مما يدل على ضعف قوة الاختبار لتلك الفروض ، وتراوحت قيمته بين (٠,٠٥ إلى ٠,٩٩) .**
- **نظرا لأن المتوسط الحسابى يتأثر بالقيم المتطرفة ، قام الباحث بحساب عدد الفروض التى بلغت قيمتها أعلى من (٠,٨٠) و أقل من (٠,٨٠) ، بالنسبة للفروض التى رفضت الفرض الصفرى بلغ عدد الفروض التى قوتها أعلى من (٠,٨٠) حوالى (١٤٤) فرضا بنسبة حوالى (٦٣,١٥٪) بينما بلغت عدد الفروض التى أقل من (٠,٨٠) حوالى (٢١) فرضا بنسبة بلغت (٩,٢١٪) ، أما الفروض التى قبلت الفرض الصفرى فقد بلغ عدد الفروض التى قوتها الإحصائية أقل من (٠,٨٠) حوالى (٦١) فرضا بنسبة بلغت (٢٦,٧٥٪) ، أما عدد الفروض التى أعلى من (٠,٨٠) فبلغت حوالى (٢) بنسبة بلغت (٠,٨٧٪).**

- المتوسط العام لقوة الاختبار للفروض ككل بلغ (٠,٧٢) بمدى يتراوح بين (٠,٠٥ إلى١,٠٠) ، حيث بلغ عدد الفروض التى أقل من (٠,٨٠) حوالى (٨٢) بنسبة بلغت (٣٥,٩٦٪) أما الفرض التى أعلى من (٠,٨٠) فبلغ عددها (١٤٦) بنسبة بلغت (٦٤,٠٢٪) .

وقد اختلفت تلك النتائج مع دراسة (فيصل أحمد عبد الفتاح ، ٢٠١١) حيث توصلت إلى أن عدد الفروض التى فشلت فى رفض الفرض الصفرى حوالى (٤٣,٩٧٪) والفروض التى رفضت الفرض الصفرى حوالى (٥٦,١٪) ، كما بلغت نسبة (٧١,٦٪) من الفروض القيمة (٠,٨٠) أو زادت عنها ، وهذا يدلل على أن حوالى ربع الفروض التى تم رفضها لم تصل إلى مستوى قوة الاختبار الإحصائي بينما بلغت نسبة الفروض التى رفضت الفرض الصفرى ووصلت لقوة الاختبار حوالى (٨٧,٢٧٪) بينما الفروض التى لم تصل لقوة الاختبار الإحصائي بلغت (١٢,٧٢٪) ، وهى قيمة أقل من قيمة دراسة (فيصل أحمد عبد الفتاح ، ٢٠١١) لكن هذا يجعل الباحثين يشككون فى نتائج حوالى (١٢,٧٢٪) من الفروض .

للاجابة عن السؤال الثانى والذى ينص على " ما واقع استخدام حجم التأثير فى الأبحاث المنشورة بمجلة التربية وعلم النفس بالمنيا "

تم تحليل الدراسات المنشورة بالمجلة منذ تاريخ (٢٠١٠ وحتى نهاية ٢٠١٢) وبلغت(٩٠) دراسة وجد الباحث منها (٦٠) دراسة استخدمت أساليب إحصائية يمكن استخدام حجم التأثير معها من تلك الأساليب: (معامل الارتباط، اختبار "ت"، تحليل التباين ، تحليل الانحدار، مان ويتنى، وويلككسون، ومربع كاى والنسبة الحرجة).

قام الباحث بحساب نسب استخدام حجم التأثير مع تلك الأساليب فجاءت النتائج كما يتضح من الجدول (٥) :

جدول (٥) : نسب استخدام حجم التأثير مع الأساليب الإحصائية

الإحصاء	معامل الارتباط	اختبارت	تحليل التباين	تحليل الانحدار	الإحصاء اللابارامترى	المتوسط النسبى
عدد الدراسات التي استخدمت حجم التأثير مع الأسلوب	صفر	١٢	٣	١	٢	
نسبة الاستخدام	صفر	٢٠٪	٥٪	١,٦٪	٣,٣٪	٦٪

يتضح من الجدول (٥) ما يلى :

- عدد الدراسات التى استخدمت حجم التأثير كمكمل للدلالة الإحصائية بلغت حوالى (صفر) لمعامل الارتباط، و(٢٠٪) لاختبار "ت"، و(٥٪) لتحليل التباين ، و (١,٦٪) لتحليل الانحدار، و(٣,٣٪) للإحصاء اللابارامترى، وبلغ المتوسط الوزنى حوالى (٦٪) .
- ضعف استخدام حجم التأثير كمكمل إحصائى للدلالة الإحصائية فى البحوث التربوية والنفسية فى المجلة .

بالنسبة للسؤال الثالث : ما واقع حجم التأثير للفروض الدالة وغير الدالة إحصائيا لاختبارى "ت" وتحليل التباين ؟

للإجابة عنه تم اشتقاق سؤالين منه هما كما يلى :
ما واقع حجم التأثير للفروض الدالة وغير الدالة إحصائيا لاختبارى "ت"؟

تم تطبيق ذلك على اختبار ت للدراسات التى تضمنت جداولها بيانات يمكن من خلالها حساب حجم التاثير وذلك لان اختبار ت كان اكثر الاختبارات تكرارا حيث استخدمته حوالي (٤٦) دراسة وتم حصر عدد الفروض الدالة احصائيا فبلغت (٢٠٦) فرض بينما بلغ عدد الفروض غير الدالة احصائيا والتى تقدم جداولها بيانات يمكن حساب حجم التأثير من خلالها (٨٤) فرضا، تم حساب حجم التأثير بمربع ايتا وكانت النسب كما فى الجدول (٦) :

جدول (٦) : مستويات حجم التأثير للفروض الدالة وغير الدالة باختبار "ت"

الدلالة	الفروض	حجم التأثير ضعيف	حجم التأثير متوسط	حجم التأثير كبير
دال	العدد	٣٦	٢٣	١٤٧
	النسبة	١٧٪	١١٪	٧٢٪
غير دال	العدد	٨٠	١	٣
	النسبة	٩٥,٢٪	١,٢	٣,٦٪

يتضح من الجدول (٦) ما يلى :

- أن نسبة (١٧٪) من الفروض الدالة إحصائيا كان حجم التأثير لها ضعيفا ونسبة (١١٪) كان متوسطا ، ونسبة (٧٢٪) كان قويا ، أى أن حوالى (٢٨٪) من الفروض الدالة إحصائيا كان حجم التأثير لها يتراوح بين الضعيف والمتوسط .
- أن نسبة (٩٥,٢٪) من الفروض غير الدالة كان حجم التأثير لها ضعيفا ونسبة (١,٢٪) متوسطا ، وحوالى (٣,٦٪) قويا ،وهذا يوضح العلاقة بين مستوى الدلالة وحجم التأثير لكن النسبة (٣,٦٪) من الفروض ليست بالقليلة لذا يوصى البحث باستخدام حجم التأثير فى كل الأحوال سواء الفرض دال أو غير دال إحصائيا .
- ما واقع حجم التأثير للفروض الدالة وغير الدالة احصائيا باختبار تحليل التباين ؟

تم تحليل الفروض الدالة إحصائيا والتى استخدمت أسلوب تحليل التباين حيث بلغت الفروض المستخدمة فى الدراسات والدالة إحصائيا ومقدمة بيانات يمكن حساب حجم التأثير من خلالها (٨٢) فرضا بينما بلغ عدد الفروض غير الدالة والتى جداولها تقدم بيانات يمكن من خلالها حساب حجم التأثير (٩٠) فرضا ، تم اجراء حساب حجم التاثير لها بمربع ايتا ثم تقدير حجم التأثير ضعيف ، متوسط ، كبير كما يتضح من الجدول (٧) :

جدول (٧) : مستويات حجم التأثير للفروض الدالة وغير الدالة باختبار تحليل التباين

الدلالة	الفروض	حجم التأثير ضعيف	حجم التأثير متوسط	حجم التأثير كبير
دال	العدد	٦٤	٩	٩
	النسبة	٧٨٪	١١٪	١١٪
غير دال	العدد	٩٠	–	–
	النسبة	١٠٠٪	–	–

يتضح من الجدول ما يلى :

- بالنسبة للفروض الدالة إحصائيا بلغت نسبة الفروض ذات حجم التأثير الضعيف (٧٨٪)، بينما بلغت نسبة الفروض ذات حجم التأثير المتوسط (١١٪)، والفروض ذات التأثير الكبير (١١٪) .
- بالنسبة للفروض غير الدالة إحصائيا بلغت نسبة الفروض ذات حجم التأثير الضعيف (١٠٠٪).
- يتضح من النتائج أن حجم التأثير مكمل إحصائى للدلالة الإحصائية إذا كان الفرض دالا ، ولكن ليست هناك ضرورة لاستخدام حجم التأثير مع الفروض غير الدالة .

السؤال الرابع :هل توجد علاقة ارتباطية دالة بين حجم التأثير وقوة الاختبار الإحصائي ؟

للتحقق من ذلك قام الباحث باستخدام معامل ارتباط سبيرمان لايجاد العلاقة بين حجم التأثير وقوة الاختبار الاحصائى حيث قام الباحث يدويا بحساب حجم التأثير (d) لكوهين - وذلك لضرورة استخدامه فى حساب قوة الاختبار الإحصائي – للفروض التى استخدمت اختبار ت نظرا لانه أكثر الفروض تكرارا وقد جاءت النتائج كما يتضح من الجدول (٨)

جدول (٨) : العلاقة الارتباطية بين حجم التأثير وقوة الاختبار الإحصائي

	عدد الفروض	معامل الارتباط	مستوى الدلالة	مربع معامل الارتباط
الدالة	١٦٥	٠,٨٧	٠,٠١	٠,٧٦
غير الدالة	٦٣	٠,٩٦	٠,٠١	٠,٩٢
ككل	٢٢٨	٠,٩٤	٠,٠١	٠,٨٨

يتضح من الجدول (٨) ما يلى :

- معامل الارتباط للفروض الدالة (التى رفضت الفرض الصفرى) بين حجم التأثير وقوة الاختبار الإحصائى بلغ (٠,٨٧) وهى قيمة توضح وجود علاقة ارتباطية دالة موجبة بين حجم التأثير وقوة الاختبار الإحصائي حيث جاءت القيمة دالة عند مستوى (٠,٠١) ، وبلغت نسبة التباين المفسر (٠,٧٦).
- معامل الارتباط للفروض غير الدالة (التى قبلت الفرض الصفرى) بين حجم التأثير وقوة الاختبار الإحصائى بلغ (٠,٩٦) ،مما يدل على وجود علاقة ارتباطية دالة موجبة بين حجم التأثير وقوة الاختبار الإحصائي حيث جاءت قيمة معامل الارتباط دالة عند مستوى (٠,٠١)، وبلغت نسبة التباين المفسر (٠,٩٢).
- معامل الارتباط للفروض ككل بين حجم التأثير وقوة الاختبار الإحصائي بلغ (٠,٩٤) ، وهو يوضح وجود علاقة ارتباطية دالة وموجبة عند مستوى (٠,٠١) بين حجم التأثير وقوة الاختبار الإحصائي ، وبلغت نسبة التباين المفسر (٠,٨٨) .

السؤال الخامس : هل توجد علاقة ارتباطية دالة بين مستوى الدلالة وكل من قوة الاختبار الإحصائى وحجم التأثير ؟

للإجابة عن السؤال تم استخدام معامل الارتباط بين مستوى الدلالة وحجم التأثير ، ثم ايجاد العلاقة الارتباطية بين مستوى الدلالة وقوة الاختبار

الإحصائي ثم ايجاد مربع معامل الارتباط لايضاح نسبة التباين المفسر والجدول (٨) يوضح ذلك .

جدول (٩) : العلاقة الارتباطية بين مستوى الدلالة وكل من قوة الاختبار الإحصائى وحجم التأثير

	عدد الفروض	معامل الارتباط	مستوى الدلالة	مربع معامل الارتباط
قوة الاختبار	٢٢٨	٠,٧٨	٠,٠١	٠,٦١
حجم التأثير	٢٢٨	٠,٧٠	٠,٠١	٠,٤٩

يتضح من الجدول (٩) ما يلى :

- وجود علاقة ارتباطيه دالة وموجبة بين مستوى الدلالة وقوة الاختبار الإحصائي حيث بلغت قيمة معامل الارتباط (٠,٧٨) ، وذات نسبة تباين مفسر بلغت (٠,٦١).
- وجود علاقة ارتباطيه دالة وموجبة بين مستوى الدلالة وحجم التأثير حيث بلغت قيمة معامل الارتباط (٠,٧٠) ، وبلغت نسبة التباين المفسر (٠,٤٩).

- **التفسير والاستنتاجات** :

يتضح من النتائج أن الدلالة الإحصائية وحدها لاتكفى كدليل على رفض الفرض الصفرى أو قبوله، حيث من المفيد أن يستخدم الباحث الدلالة العملية (حجم التأثير)، وقوة الاختبار الإحصائي لتجنب الوقوع فى الخطأ من النوع الأول ، والخطأ من النوع الثانى .

حيث أوضحت النتائج انخفاض متوسط قوة الاختبار الإحصائي للفروض ككل للأعداد المنشورة فى مجلة البحث فى التربية وعلم النفس منذ عام ٢٠١٠ وحتى نهاية ٢٠١٢ ، ومن ثم فإنه يوجد تشكك فى نتائج تلك الفروض وخاصة الفروض التى رفضت الفرض الصفرى ، ولذا يناشد البحث القائمين على المجلات العلمية فى المجتمع العربي بضرورة الالتزام بدراسة قوة الاختبار الإحصائي وتحديدها مسبقا قبل النشر فى المجلة ، وإلزام الباحثين بذلك حتى يكون القرار المأخوذ عن نتائج الأساليب الإحصائية قرارا صائبا .

كما أوضحت النتائج ضعف استخدام حجم التأثير كمكمل إحصائي للدلالة الإحصائية حيث بلغ المتوسط الوزنى للدراسات التى استخدمت حجم التأثير (٦٪) فقط ، وهى نسبة تشير إلى أن أكثر من (٩٤٪) من البحوث لم تستخدم حجم التأثير ، كما أن نسبة (٢٨٪) من الفروض الدالة التى استخدم معها اختبار "ت" كانت دلالتها العملية تتراوح بين ضعيف ومتوسط ، بينما حوالى (٨٩٪) من الفروض الدالة إحصائيا والتى اعتمد فى التحقق منها على تحليل التباين كانت دلالتها العملية تتراوح بين ضعيف ومتوسط ، وهذا يدعو إلى ضرورة إلزام الناشرين بالمجلات العلمية بضرورة وضع الدلالة العملية مع الدلالة الإحصائية فى بحوثهم كمتطلب أساسي للنشر حتى تتضح القيمة العملية للبحث المنشور .وتتفق تلك النتائج مع ما توصلت إليه دراسة (Hasse et.al.,1982) ، و دراسة (عبدالعاطي أحمد الصياد ، ١٩٨٨) ، و دراسة (عبد الله عمر النجار، ١٩٩٠)، ودراسة ابتسام حسن الصائغ (١٩٩٦) ، و دراسة إبراهيم رشاد (٢٠٠٠) ، ودراسة عادل أحمد بابطين (٢٠٠١) .

وبالنسبة للفروض غير الدالة ،جاءت نسبة (١٠٠٪) من الفروض ذات حجم تأثير ضعيف فى حال الفروض التى اعتمد فى التحقق منها على تحليل التباين ، بينما بلغت نسبة (٣,٦٪) من الفروض ذات حجم تأثير قوى للفروض التى تم التحقق منها باستخدام اختبار "ت" ، لذا يوصى البحث بضرورة وضع الدلالة العملية مع الدلالة الإحصائية فى حال الفروض التى تم التحقق منها باستخدام اختبار "ت" لكل من الفروض الدالة وغير الدالة إحصائيا، بينما ليست هناك ضرورة لذلك فى حال الفروض غير الدالة إحصائيا والتى اعتمدت على تحليل التباين .وتتفق النتائج لحد ما مع نتائج دراسة تحسين حجيمات ، و خليل عليان (١٩٩٧) ، ودراسة (Mclean & Eronest,1997) ، و دراسة ألطاف إبراهيم (١٩٩٨)

وبالنسبة للعلاقة الارتباطية بين قوة الاختبار الإحصائي وكل من حجم التأثير ومستوى الدلالة ، فقد أوضحت النتائج وجود علاقة ارتباطية دالة موجبة بين حجم التأثير وقوة الاختبار الإحصائي، مما يوضح أنه كلما زاد حجم التأثير زادت قوة الاختبار الإحصائى ، كما يوضح البحث وجود علاقة ارتباطية دالة موجبة بين حجم التأثير ومستوى الدلالة فقوة الاختبار الإحصائي للاختبارات الدالة إحصائيا أعلى من قوة الاختبار الإحصائي للفروض غير الدالة ، والاختبار الدال عند مستوى (٠,٠١) تكون قوته الإحصائية أعلى من الاختبار الدال عند مستوى (٠,٠٥) .

وتتفق تلك النتائج مع ما توصلت إليه دراسة (عبد الله عمر النجار ،٢٠٠٦) من أن احتمال وجود الخطأ من النوع الثانى (أي احتمال وجود قوة الاختبار) منخفضة مع حجم التأثير المنخفض أكبر من احتمال وجود هذا الاقتران مع حجم التأثير المتوسط والعالى ، وأن قوة الاختبار الإحصائي فى المتوسط عالية نظرا لزيادة حجم العينة وليس لأن الباحثين تقصدوا الاهتمام بقوة الاختبار وأن هذه الزيادة ربما كانت وسيلة تخفى وراءها الرغبة فى الحصول على دلالة إحصائية لدلالة عملية ضعيفة ، كما أن زيادة حجم العينة ينقص من قيمة الخطأ من النوع الأول وبذلك تتأثر قوة الاختبار نظرا للعلاقة الطردية بينهما.

ومن ثم يخرج البحث بعدة جمل مفيدة تتمثل فى : أن الدلالة العملية هى وجه مكمل للدلالة الإحصائية ، وأنه يجب على الباحثين مراعاة قوة الاختبار الإحصائي عند تحديد حجم العينة ، وفرض الفروض ، وتحديد مستوى الدلالة ، تجنبا للوقوع فى الخطأ من النوع الثانى ، وحتى يصبح القرار المتخذ برفض الفرض الصفرى قرارا صائبا يمكن الاستناد عليه فيما بعد ، كما يجب على الباحثين تضمين البحث بيانات يمكن من خلالها تحدة القوة الإحصائية للنتائج التى تم التوصل إليها .

- **المراجع** :
- **أولا** : **المراجع العربية** :

- ابتسام حسن الصائغ (١٩٩٦) .الدلالة الإحصائية والدلالة العملية في اختبار(ت) و(ف) :دراسة تقويمية مقارنة لرسائل الماجستير في كلية التربية بجامعة أم القرى ،رسالة ماجستير غير منشورة ، كلية التربية، جامعة أم القرى بمكة المكرمة.

- إبراهيم رشاد (٢٠٠٠) .واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبارللاختبارات الإحصائية المستخدمة في بحوث مجلة دراسات العلوم التربوية في الأعوام (١٩٩٧- ١٩٩٨)، في الجامعة الأردنية.رسالة ماجستير غير منشورة، الأردن الجامعة الأردنية .
- أحمد سليمان عوده، و خليل يوسف الخليلي (٢٠٠٠) .الإحصاء للباحث في التربية والعلوم الإنسانية، إربد، دار الأمل.
- السيد محمد أبو هاشم (٢٠٠٥). مؤشرات التحليل البعدىMeta- Analysis لبحوث فعالية الذات فى ضوء نظرية باندورا .مركز بحوث التربية (جامعة الملك سعود)، ٢٣٨ ، ١- ٩٠ .
- ألطاف إبراهيم (١٩٩٨) .واقع الدلالة الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في رسائل ماجستير علم النفس التربوي في الجامعة الأردنية . رسالة ماجستير غير منشورة ، الأردن ، الجامعة الأردنية.
- تحسين حجيمات، و خليل عليان (١٩٩٧) . واقع الدلالة الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في رسائل ماجستير الارشاد النفسي والتربوي بالجامعة الأردنية . مجلة العلوم التربوية ، ٢٤ (٢) ، ٣٩٨- ٤٠٨
- خليل أحمد محمود لبد(٢٠٠٥). تقوىم بعض الإجراءات المن□جىة المستخدمة في رسائل الماجستىر المقدمة لكلىات الترب‌ىة في الجامعات الفلسطىنىة بغزة ،رسالة ماجستير غير منشورة، كلية التربية ، الجامعة الإسلامية – غزة .
- رجاء محمود أبوعلام (٢٠٠٧) .مناهج البحث في العلوم النفسية والتربوية ،القاهرة :دار النشر للجامعات.
- رشدى فام (١٩٩٧) . حجم التأثير – الوجه المكمل للدلالة الإحصائية . المجلة المصرية للدراسات النفسية ، ٧ (١٦) ، ٥٧ – ٧٥ .
- زكريا الشربينى (٢٠٠٧). الإحصاء وتصميم التجارب فى البحوث النفسية والتربوية والاجتماعية ، القاهرة ، الأنجلو المصرية .
- زكريا الشربينى (٢٠٠٧). الإحصاء اللابارامترى فى العلوم النفسية والتربوية ، القاهرة ، الأنجلو المصرية .
- طلال هيازع حسن البارقى (٢٠١٢).واقع الدلالة الإحصائية والدلالة العملية للبحوث المنشورة بمجلة جامعة أم القرى للعلوم التربوية والاجتماعية والإنسانية فى المدة من ١٤٢٥- ١٤٣٠ ه ، رسالة ماجستير غير منشورة، كلية التربية ، جامعة أم القرى .
- عبدالعاطي أحمد الصياد (١٩٨٨) .الدلالة العملية وحجم العينة المصاحبتين للدلالة الإحصائية لاختبار–ت –في البحث التربوي والنفسي العربي ،بحوث مؤتمر البحث التربوي بين الواقع والمستقبل ،رابطة التربية الحديثة ، القاهرة ،٢ - ٤ يوليو ، ١٩٧- ٢٣٣.
- عبد الله عمر النجار (١٩٩١) .دراسة تقويمية مقارنة للأساليب الإحصائية التي استخدمت في تحليل البيانات رسائل الماجستير في كل من كلية التربية بجامعة أم القرى بمكة المكرمة وكلية التربية بجامعة الملك سعود بالرياض ، رسالة ماجستير غير منشورة ، جامعة أم القرى.

- عبد الله عمر النجار (٢٠٠٥). الدلالة الإحصائية والدلالة العليمة لاختبار كا٢ فى البحوث الإدارية المنشورة (دراسة تقويمية) . المجلة العربية للعلوم الإدارية . ١٢ (٢) .١٦٩- ١٩٣

- عبد الله عمر النجار(٢٠٠٦).دراسة تحليلية لقوة الاختبار الإحصائى فى البحوث الإدارية المنشورة . المجلة العلمية لجامعة الملك فيصل (العلوم الإنسانية والإدارية) . ٧ (٢) ، ٢٦١ - ٢٩٣

- عبد الله المنيزل (٢٠٠٠). الإحصاء الاستدلالى وتطبيقاته فى الحاسوب باستخدام الرزم الإحصائية Spss، عمان ، دار وائل للنشر .

- عادل أحمد بابطين (٢٠٠١) .مشكلات الدلالة الإحصائية في البحث التربوي وحلول بديلة، رسالة ماجستير غير منشورة ، كلية التربية، جامعة أم القرى.

- فريال محمود الحاج (٢٠٠٣) .واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار للاختبارات الإحصائية المستخدمة في رسائل الماجستير الصادرة عن كلية العلوم التربوية. رسالة ماجستير غير منشورة : جامعة مؤتة.

- فريد أبو زينه ، و عبد الحافظ الشايب ، وعماد عباينة ، ومحمد عبد العال التعيمى (٢٠٠٧) . مناهج البحث العلمى والإحصاء فى البحث العلمى . عمان ، دار المسيرة .

- فيصل أحمد عبد الفتاح (٢٠١١).قوة الاختبارات الإحصائية وحجم الأثر فى البحوث التربوية والنفسية ، مجلة العلوم التربوية والدراسات الإسلامية (٢)، جامعة الملك سعود م٢٣ ، ١٩٥- ٢١٤ .

- فيصل أحمد العبد القادر (٢٠٠٩).حجم تأثير الاختبارات المعملية واللامعملية المستخدمة فى رسائل الماجستير بكلية التربية بجامعة الملك سعود ، رسالة ماجستير غير منشورة كلية التربية ، جامعة الملك سعود .

- محمد إبراهيم أحمد الشاردى (١٤٣٣هـ).تأثير حجم العينة على قوة الاختبار الإحصائي .رسالة ماجستير غير منشورة ، كلية التربية ، جامعة أم القرى .

- مصطفى حسين باهى (٢٠١١).بعض القضايا الهامة للتحليل الإحصائي فى البحوث النفسية والتربوية ، المؤتمر السنوى العربى السادس –الدولى الثالث "تطوير برامج التعليم العالى النوعى فى مصر والوطن العربي فى ضوء متطلبات عصر المعرفة "، كلية التربية النوعية ، جامعة المنصورة ، ١٣- ١٤ أبريل ، ١٣٨٧- ١٤١٤

- يحي نصار (٢٠٠٢).حجم الأثر كأسلوب إحصائي مكمل لفحص الفرضيات الإحصائية . مركز بحوث كلية التربية (جامعة الملك سعود) ، ١٧٦ ، ١- ٤١ .

- يحي نصار (٢٠٠٦).استخدام حجم الأثر لفحص الدلالة العملية للنتائج فى الدراسات الكمية ، مجلة العلوم التربوية والنفسية ، كلية التربية ، جامعة البحرين ، ٧ (٢)، ٣٥- ٦٠

- **ثانيا- المراجع الأجنبية** :

- Cashen, L.H. & Geiger , S.W. (2004). Statistical power and the testing of null hypotheses : A review of contemporary management research and recommendations for future studies. *Organizational Research Methods ,7,* 151-167.doi:10.1177/1094428104263676

- Cohen, J. (1962). The statistical power of abnormal–social psychological research: A review. Journal of Abnormal Psychology, 65, 145–153.
- Cohen, J. (1977). Statistical power analysis for the behavioral sciences, New York: Academic Press.
- Cohen. J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1992). A power primer. Psychological Bulletin, 112, 155-159.
- Cohen,J.(1992).Statistical Power Ananlysis . Current Direction in Psychological Science , 1(3) , 98-101.
- Cohen, J. (1994). The earth is round (p < 0.05). American Psychologist, 49, 997-1003.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191.
- Haase, R. F., Waechter, D. M., & Solomon, G. S. (1982). How significant is a significant difference? Average effect size of research in counseling psychology, Journal of Counseling Psychology ,29 , 58-65.
- Hallahan, Mark and Rosenthal , Robert(1996). Statistical power: concepts, procedures, and applications. Behav. Res. Ther. Vol. 34, No. 5/6, 489-499.
- Howell, D.C.(1995).Fundamental statistics for the behavioral sciences (3rd ed.) .CA: Belmont.
- Huston, Holly L.(1993). Meaningfulness, Statistical Significance, Effect Size,and Power Analysis: A General Discussion with Implications for MANOVA. paper presented at annual meeting of the mid-south educational research association (2nd, new orleans, la, november 9-12, 1993).
- Kellow, J. Thomas.(1998). Beyond Statistical Significant Tests: The Importance of Using Other Estimates of Treatment Effects To Interpret Evaluation Results. American Journal of Evaluation. V19 n1, P123-34, Win 1998.
- Maxwell, S. E., & Delaney, H. D. (1990). Designing experiments and analyzing data: A model comparison perspective. Belmont, CA: Wadsworth.
- Mcclain, A.(1995).Effect size as an alternative to statistical significance testing. Paper presented at the annual meeting for the

American Educational Research Association. San Francisco,CA.(Eric Document Reproduction Service No. ED382639).

- Mclean, J.E, Eronest,J.M.(1997). Has testing for statistical significance outlived its usefullness,.Paper presented at the the annual meeting of the Mid-south Educational Research Association (26,Memphis, TN, November 12-14, 1997). (Eric Reproduction ServiceNo.Ed415265).
- McLean, James E., & Kaufman, Alan S.(1998a). Statistical Significance Testing. Research in the Schools. V5 n2, Fall 1998.
- McLean, James, & Ernest, James M.(1998b).The Role of Statistical Significance Testing in Educational Research. Research in the Schools.V5 n2,P15-22,Fall 1998.
- Nix, T. W., & Barnette, J. J. (1998). The data analysis dilemma: ban or abandon. A review of null hypothesis significance testing. Research in the Schools, 5(2)3-14.
- Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. Contemporary Educational Psychology, 25, 241–286.
- Parker,R and Hagan-Burke, S(2007).Useful Effect Size Interpretations for Single Case Research . Behavior Therapy , 83 , 95-105.
- Portillo ,Matthew T. (2001).Statistical Significance Tests and Effect Magnitude Measures, within quantitative research manuscripts published in the journal of agricultural education during 1996-2000.
- Shaver, James P.(1992). What Statistical Significance Testing Is, and What It Is Not. Paper Presented at Annual Meeting of the American Educational Research Association (San Francisco, CA,April 20-24, 1992).
- Snyder, P., & Lawson , S. (1993). Evaluating results using corrected and uncorrected effect size estimates. Journal of Experimental Education ,61, 334- 349.
- Steinberg, L and Thissen , D (2006).Using Effect Sizes for Research Reporting : Examples Using Item Response Theory to Analyze differential item functioning . Psychological Methods , 11,4,402-415.

****